

Free energy and complexity of spherical bipartite models

Antonio Auffinger ^{*}

University of Chicago

Wei-Kuo Chen [†]

University of Chicago

May 12, 2014

Abstract

We investigate both free energy and complexity of the spherical bipartite spin glass model. We first prove a variational formula in high temperature for the limiting free energy based on the well-known Crisanti-Sommers representation of the mixed p -spin spherical model. Next, we show that the mean number of local minima at low levels of energy is exponentially large in the size of the system and we derive a bound on the location of the ground state energy.

1 Introduction and main results

In recent years, spin glass theory has been successfully extended beyond the area of mathematics and mathematical physics, for instance, to the study of social and neural networks [4, 10, 17, 21] or theoretical biology and computer science [2, 8, 9, 19]. Motivated by these fields of research, bipartite spin systems were introduced to model collective properties of two (or more) interactive large groups of individuals. Although very attractive for their realm of applications, bipartite spin glasses are believed to be very difficult to analyze even non-rigorously. For instance, a Parisi-like theory for these models is still under construction within the physics community, and the models lack many of the predictions that other mean field spin glasses share, e.g., the Sherrington-Kirkpatrick (SK) model. A recent account of the predictions can be found in [11] and the references therein.

In this paper, we start a rigorous treatment on one class of disordered bipartite spin systems, known as the mixed (p, q) -spin spherical bipartite models. Our results concern about the behaviors of the free energy and the extreme values of the Hamiltonian. First, we obtain a variational formula for the limiting free energy at very high temperature in terms of the well-known Crisanti-Sommers formula for the mixed p -spin spherical model [22, 15]. Surprisingly, our expression (see Theorem 1) is quite different from what has been expected in the bipartite SK model with Ising spin, where the limiting free energy at high temperature is conjectured to be a minimax formula proposed in [11, 12]. Second, we provide estimates for the complexity of local minima at any level of energy. We give bounds on the mean number

^{*}auffing@math.uchicago.edu

[†]wkchen@math.uchicago.edu

of local minima of the Hamiltonian whose values are below any threshold. These estimates lead to a bound on the location of the ground state energy and establish that the system has a positive complexity. In the physics literature, models with positive complexity are widely known to show glass to spin-glass phase transitions for instance. However, rigorous results about the complexity are scarce. We believe that our results are the first evidence that the bipartite models share the same phenomena.

We now introduce the spherical bipartite model as follows. For a positive integer M , denote by S^M the $(M - 1)$ -sphere with radius $M^{1/2}$ and by λ_M the uniform probability measure on S^M . For any $x = (x_1, \dots, x_M)$ and $y = (y_1, \dots, y_M)$ in \mathbb{R}^M , set

$$R_M(x, y) = \frac{1}{M} \sum_{i=1}^M x_i y_i.$$

Let N, N_1, N_2 be positive integers that satisfy $N = N_1 + N_2$ and

$$\left| \frac{N_1}{N} - \gamma \right| = o(N) \quad (1)$$

for some $\gamma \in (0, 1)$. For any $p, q \geq 1$, we set the pure (p, q) -spin bipartite Hamiltonian as

$$H_{N,p,q}(u, v) = \sum_{1 \leq i_1, \dots, i_p \leq N_1} \sum_{1 \leq j_1, \dots, j_q \leq N_2} g_{i_1, \dots, i_p, j_1, \dots, j_q} u_{i_1} \dots u_{i_p} v_{j_1} \dots v_{j_q}$$

for $u = (u_1, \dots, u_{N_1}) \in S^{N_1}$ and $v = (v_1, \dots, v_{N_2}) \in S^{N_2}$, where all g variables are centered i.i.d. Gaussian random variables with variance $N/(N_1^p N_2^q)$. A computation reveals that

$$\mathbb{E} H_{N,p,q}(u, v) H_{N,p,q}(u', v') = N (R_{N_1}(u, u'))^p (R_{N_2}(v, v'))^q.$$

The Hamiltonian of the mixed-spin bipartite spherical model is defined as

$$H_N(u, v) + h_1 \sum_{i=1}^{N_1} u_i + h_2 \sum_{i=1}^{N_2} v_i,$$

where H_N is the linear combination of $H_{N,p,q}$,

$$H_N(u, v) = \sum_{p,q \geq 1} \beta_{p,q} H_{N,p,q}(u, v).$$

Here $h_1, h_2 \in \mathbb{R}$ denote the strength of the external fields and the nonnegative sequence $(\beta_{p,q})_{p,q \geq 1}$ is usually called the (inverse) temperature parameter that satisfies $\beta_{p,q} > 0$ for at least one pair (p, q) and decreases fast enough, for instance, $\sum_{p,q=1}^{\infty} 2^{p+q} \beta_{p,q}^2 < \infty$. This assumption ensures the well-definedness of the model and a direct computation gives

$$\mathbb{E} H_N(u, v) H_N(u', v') = N \xi(R_{N_1}(u, u'), R_{N_2}(v, v'))$$

for

$$\xi(x, y) := \sum_{p,q \geq 1} \beta_{p,q}^2 x^p y^q, \quad \forall x, y \in [0, 1]. \quad (2)$$

1.1 Free energy

We first devote our attention to understanding the bipartite spherical model under the assumptions that the temperature parameter $\xi(1, 1)$ and the external fields h_1, h_2 are sufficiently small. Define the Gibbs measure of the bipartite spherical model on $S^{N_1} \times S^{N_2}$ by

$$dG_N(u, v) = \frac{1}{Z_N} \exp \left(H_N(u, v) + h_1 \sum_{i=1}^{N_1} u_i + h_2 \sum_{i=1}^{N_2} v_i \right) d(\lambda_{N_1} \times \lambda_{N_2})(u, v),$$

where the normalizing factor Z_N is called the partition function. For a fixed realization g , we denote by (u^1, v^2) and (u^2, v^2) i.i.d. sampled configurations (replicas) with respect to G_N . The overlaps within the first and second parties are defined respectively as

$$R_{1,2}^1 = R_{N_1}(u^1, u^2) = \frac{1}{N_1} \sum_{i=1}^{N_1} u_i^1 u_i^2 \text{ and } R_{1,2}^2 = R_{N_2}(v^1, v^2) = \frac{1}{N_2} \sum_{i=1}^{N_2} v_i^1 v_i^2.$$

Finally, we set the free energy as

$$F_N = \frac{1}{N} \mathbb{E} \log Z_N.$$

First, we will show that the limiting free energy can be computed through an analogy of the well-known Cransti-Sommers formula in the mixed p -spin spherical model. Second, we will show that the overlap within each party is concentrated around a constant with a qualitative error estimate. For notational convenience, throughout this section we will use $M_0, \lambda_0 > 0$ to stand for very small quantities and L is a universal constant independent of everything. These quantities might be different from each occurrence. Also, as we will only be concerned with very small $\xi(1, 1) = \sum_{p,q \geq 1} \beta_{p,q}^2$, we will use the clear fact without mentioning at each time that the partial derivatives of ξ up to the second order are uniformly small as well.

We now state our main results. Define $L : [0, 1]^2 \rightarrow \mathbb{R}$ by

$$\begin{aligned} P(a, b) = & \frac{\gamma}{2} \left(h_1^2(1-a) + \frac{a}{1-a} + \log(1-a) + \xi(1, 1) - \xi(a, b) \right) \\ & + \frac{1-\gamma}{2} \left(h_2^2(1-b) + \frac{b}{1-b} + \log(1-b) + \xi(1, 1) - \xi(a, b) \right). \end{aligned}$$

Theorem 1. *There exists a small $M_0 > 0$ such that the following statements hold whenever $\xi(1, 1), |h_1|, |h_2| < M_0$.*

(i) *We have that*

$$\lim_{N \rightarrow \infty} F_N = \min_{a, b \in [0, 1]} P(a, b). \quad (3)$$

Here, the minimum is attained by $(a_0, b_0) \in [0, 1]^2$ that is the unique solution to the following system of equations

$$h_1^2 + \frac{\partial_x \xi(a_0, b_0)}{\gamma} = \frac{a_0}{(1-a_0)^2}, \quad h_2^2 + \frac{\partial_y \xi(a_0, b_0)}{1-\gamma} = \frac{b_0}{(1-b_0)^2}.$$

(ii) Let $\langle \cdot \rangle$ denote the expectation with respect to the product measure $G_N \times G_N$. For the overlaps,

$$\mathbb{E} \langle (R_{1,2}^1 - a_0)^2 \rangle, \mathbb{E} \langle (R_{1,2}^2 - b_0)^2 \rangle \leq \frac{L \log^2 N}{N}, \quad (4)$$

where L is a universal constant independent of everything.

Theorem 1 is the high temperature results ($\xi(1, 1)$ very small). In [11, 12], the limiting free energy of the bipartite SK model in the Ising spin case is conjectured to be a variational formula of minimax type. The formula (3) shows that the Cransti-Sommers formula in our bipartite model does not share the minimax characterization. In the case of $h_1 = h_2 = 0$, we clearly see that the overlap within each party is concentrated at 0 sharing the same behavior as the Haar measure on the product of the two spheres.

The proof of Theorem 1 is mainly motivated by a beautiful argument of Latala as presented in [20] and Chapter 1 of [23]. His argument originated from the study of the SK model in Ising spin case in the high temperature regime. We found that it works equally well in the bipartite model. To extend Theorem 1(i) to the low temperature case, i.e., $\xi(1, 1)$ is very large, a reasonable approach is to adapt the treatments in the study of the spherical SK model [22], where the Guerra replica symmetry breaking bound (RSB) has played an essential role in the computation of the limiting free energy. However, an equivalent formulation of the Guerra RSB bound does not seem to hold for any clear reason in the setting of the bipartite model, see Remark 1. Some new ideas are needed.

1.2 Complexity of local minima

Our goal in this section is to compute the complexity of local minima of the model, i.e., the mean number of local minima below a certain level of the function H_N . This computation has been carried out in the mixed p -spin spherical model in [5, 6] and also in other examples of disordered systems (see [18] and the references therein). Note that as ξ has continuous partial derivatives up to fourth order on $[0, 1]^2$, this guarantees the differentiability of H_N for almost surely all sampled realization g so that we can study the critical points of H_N . Since in this section the temperature does not play a role, we will now assume that $\xi(1, 1) = 1$. In other words, the Gaussian process H_N has constant variance N .

For any $t \in \mathbb{R}$ and $0 \leq k \leq N - 1$, the complexity function is defined as the random number $\text{Crt}_{N,k}(t)$ of critical points of index k of the function H_N below level Nt . More precisely,

$$\text{Crt}_{N,k}(t) = \sum_{(u,v): \nabla H_N(u,v)=0} \mathbf{1}\{H_N(u,v) \leq Nt\} \mathbf{1}\{i(\nabla^2 H_N(u,v)) = k\}, \quad (5)$$

where $\mathbf{1}\{E\} = \mathbf{1}_{\{E\}}$ denotes the indicator function of the event E . In particular, $\text{Crt}_{N,0}(t)$ denotes the number of local minima below level Nt . Here ∇ and ∇^2 are the gradient and the Hessian restricted to $S^{N_1} \times S^{N_2}$ and $i(\nabla^2 H_N(u,v))$ is the number of negative eigenvalues of the Hessian $\nabla^2 H_N$, called the index of the Hessian at (u,v) . Our main result is lower and upper bounds for the mean complexity of local minima.

Theorem 2. Assume that $\xi(x, y) \neq xy$. There exist continuous functions $J, K : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$J(t) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \text{Crt}_{N,0}(t) \leq K(t). \quad (6)$$

Furthermore, $(\text{Im } J) \cap (0, \infty) \neq \emptyset$ and $\lim_{t \rightarrow -\infty} K(t) = -\infty$.

The functions J and K above are derived from the analysis of the modified Hamiltonian obtained by coupling two independent mixed p -spin spherical models together, where explicit formulas are available. These functions are defined in Section 4 (see (24) and (41)) through certain variational principles. The exact expressions of J and K do not play an important role here. However, these bounds provide essential consequences. On the one hand, the upper bound in (6) gives a straightforward bound on the location of the minimum of the Gaussian process H_N as the function K takes both negative and positive values. Indeed, a simple application of the Markov inequality combined with the theorem above leads to

Corollary 1. Let m_0 be the smallest zero of the function $K(t)$. For any $\epsilon > 0$ there exist positive constants C_1, C_2 such that

$$\mathbb{P}(\min H_N \leq N(m_0 - \epsilon)) \leq C_1 \exp(-C_2 N).$$

Proof. Just note that $\min H_N \leq Nt$ if and only if $\text{Crt}_{N,0}(t) \geq 1$. □

On the other hand, the lower bound in (6) establishes what is known in the physics community as the presence of a positive complexity. Precisely, as the function J takes positive values, Theorem 2 shows an exponentially large mean number of local minima of the energy (even at deep values of H_N). Although the result here is about the mean number, it is predicted, at least in the pure case, $\xi(x, y) = x^p y^q$, that $\text{Crt}_{N,0}(t)$ concentrates around its mean. Despite the physicists believe, this self-averaging property is, as far as we know, open in *all* disordered models of spin glasses.

The rest of the paper is organized as follows. In the next section, we provide a proof of Theorem 1. In Section 3, we study the complexity of the modified Hamiltonian mentioned above. The results of Section 3 are of independent interest and are used in the proof of Theorem 2 in Section 4.

2 Proof of Theorem 1

For arbitrary $a, b \in [0, 1]$, consider the interpolated Hamiltonian H_t for $0 \leq t \leq 1$,

$$\begin{aligned} H_t(u, v) = & \sqrt{t} H_N(u, v) + \sum_{i=1}^{N_1} u_i \left(\sqrt{1 - ty_i} \sqrt{\gamma_N^{-1} \partial_x \xi(a, b)} + h_1 \right) \\ & + \sum_{j=1}^{N_2} v_j \left(\sqrt{1 - tz_j} \sqrt{(1 - \gamma_N)^{-1} \partial_y \xi(a, b)} + h_2 \right), \end{aligned}$$

where $y_1, \dots, y_{N_1}, z_1, \dots, z_{N_2}$ are i.i.d. standard Gaussian and $\gamma_N := N_1/N$. Let G_t and Z_t be the Gibbs measure and partition function associated to the Hamiltonian H_t , respectively. We denote by $(u^\ell, v^\ell)_{\ell \geq 1}$ a sequence of i.i.d. sampled configurations (replicas) with respect to G_t and for any bounded measurable function f depending on $(u^\ell, v^\ell)_{1 \leq \ell \leq n}$ for some $n \geq 1$, we use $\langle f \rangle_t$ to stand for its Gibbs expectation, i.e.,

$$\langle f \rangle_t = \int_{(S^{N_1} \times S^{N_2})^n} f(u^1, v^1, \dots, u^n, v^n) dG_t(u^1, v^1) \times \dots \times G_t(u^n, v^n).$$

Define overlaps within the first and second parties as

$$R_{\ell, \ell'}^1 = R_{N_1}(u^\ell, u^{\ell'}) = \frac{1}{N_1} \sum_{i=1}^{N_1} u_i^\ell u_i^{\ell'} \text{ and } R_{\ell, \ell'}^2 = R_{N_2}(v^\ell, v^{\ell'}) = \frac{1}{N_2} \sum_{i=1}^{N_2} v_i^\ell v_i^{\ell'}.$$

Set the interpolated free energy

$$\phi(t) = \frac{1}{N} \mathbb{E} \log Z_t.$$

First we want to control the derivative of $\phi(t)$. Using the Gaussian integration by parts,

$$\begin{aligned} \phi'(t) &= \frac{1}{2N} \mathbb{E} \left\langle \frac{H_N(u, v)}{\sqrt{t}} - \frac{\sum_{i=1}^{N_1} u_i y_i \sqrt{\gamma_N^{-1} \partial_x \xi(a, b)} + \sum_{j=1}^{N_2} v_j z_j \sqrt{(1 - \gamma_N)^{-1} \partial_y \xi(a, b)}}{\sqrt{1-t}} \right\rangle_t \\ &= -\frac{1}{2} \mathbb{E} \langle \Delta(R_{1,2}^1, R_{1,2}^2) \rangle_t + \frac{1}{2} \Delta(1, 1), \end{aligned}$$

where

$$\begin{aligned} \Delta(x, y) &:= \xi(x, y) - (x \partial_x \xi(a, b) + y \partial_y \xi(a, b)) + \theta(a, b) \\ \theta(a, b) &:= a \partial_x \xi(a, b) + b \partial_y \xi(a, b) - \xi(a, b). \end{aligned}$$

Remark 1. As the function Δ may take both positive and negative values, it does not seem so easy to see that the function ϕ' takes only one sign. For instance, $\Delta(x, y) = (x - a)(y - b)$ in the simplest case $\xi(x, y) = xy$. Therefore, it is by no means clear whether we have Guerra's bound, $F_N = \phi(1) \leq \phi(0)$, or not.

For any $x, y \in [0, 1]$, Taylor's formula says that there exist $a', b' \in [0, 1]$ such that

$$\begin{aligned} \xi(x, y) &= \xi(a, b) + \partial_x \xi(a, b)(x - a) + \partial_y \xi(a, b)(y - b) \\ &\quad + \frac{1}{2} \left\langle \nabla^2 \xi(a', b') \begin{bmatrix} x - a \\ y - b \end{bmatrix}, \begin{bmatrix} x - a \\ y - b \end{bmatrix} \right\rangle \\ &= x \partial_x \xi(a, b) + y \partial_y \xi(a, b) - \theta(a, b) \\ &\quad + \frac{1}{2} \left\langle \nabla^2 \xi(a', b') \begin{bmatrix} x - a \\ y - b \end{bmatrix}, \begin{bmatrix} x - a \\ y - b \end{bmatrix} \right\rangle. \end{aligned}$$

Using this and letting $M = \sup_{x, y \in [0, 1]} |\nabla^2 \xi(x, y)|$ and $C(x, y) = \max\{|x - p|, |y - q|\}$ give

$$|\Delta(x, y)| \leq MC(x, y). \tag{7}$$

Therefore, to control the derivative of $\Phi(t)$, it suffices to control $\mathbb{E} \langle C(R_{1,2}^1, R_{1,2}^2) \rangle_t$. We will argue by Latala's argument. For a function $f = f(u^1, v^1, \dots, u^n, v^n)$, a direct computation using again Gaussian integration by parts gives

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \langle f \rangle_t &= N \sum_{1 \leq \ell < \ell' \leq n} \mathbb{E} \langle \Delta(R_{\ell, \ell'}^1, R_{\ell, \ell'}^2) f \rangle_t - Nn \sum_{\ell \leq n} \mathbb{E} \langle \Delta(R_{\ell, n+1}^1, R_{\ell, n+1}^2) f \rangle_t \\ &\quad + \frac{n(n+1)N}{2} \mathbb{E} \langle \Delta(R_{n+1, n+2}^1, R_{n+1, n+2}^2) f \rangle_t. \end{aligned} \quad (8)$$

For any $\ell \neq \ell'$, we define

$$C_{\ell, \ell'} = C(R_{\ell, \ell'}^1, R_{\ell, \ell'}^2).$$

Observe that from Hölder's inequality and the symmetry between replicas,

$$\begin{aligned} \mathbb{E} \langle C_{\ell, \ell'}^2 \exp \lambda N C_{1,2}^2 \rangle_t &= \sum_{k=0}^{\infty} \frac{(\lambda N)^k}{k!} \mathbb{E} \langle C_{\ell, \ell'}^2 C_{1,2}^{2k} \rangle_t \\ &\leq \sum_{k=0}^{\infty} \frac{(\lambda N)^k}{k!} \left(\mathbb{E} \langle C_{\ell, \ell'}^{2k+2} \rangle_t \right)^{1/(k+1)} \cdot \left(\mathbb{E} \langle C_{1,2}^{2k+2} \rangle_t \right)^{k/(k+1)} \\ &= \mathbb{E} \langle C_{1,2}^2 \exp \lambda N C_{1,2}^2 \rangle_t. \end{aligned}$$

Using this inequality, the equation (8) with $n = 2$ and $f = \exp \lambda N C_{1,2}^2$ and (7), we obtain

$$\frac{d}{dt} \mathbb{E} \langle \exp \lambda N C_{1,2}^2 \rangle_t \leq 8MN \mathbb{E} \langle C_{1,2}^2 \exp \lambda N C_{1,2}^2 \rangle_t,$$

which yields that for $t \leq \lambda/8M$,

$$\frac{d}{dt} \mathbb{E} \langle \exp(\lambda - 8tM) N C_{1,2}^2 \rangle_t \leq 0. \quad (9)$$

From now on, for each $N \geq 1$, we pick $(a, b) = (a_N, a_N)$ such that

$$\begin{aligned} a_N &= \mathbb{E} \langle R_{1,2}^1 \rangle_0, \\ b_N &= \mathbb{E} \langle R_{1,2}^2 \rangle_0. \end{aligned}$$

Let us remark that since $(a, b) \mapsto (\mathbb{E} \langle R_{1,2}^1 \rangle_0, \mathbb{E} \langle R_{1,2}^2 \rangle_0)$ defines a continuous function on the compact and convex space $[0, 1]^2$, this ensures the existence of such (a_N, b_N) by Brouwer's fixed point theorem. We need the following lemma.

Lemma 1. *There exist small $M_0 > 0$ and $\lambda_0 > 0$ such that whenever $\xi(1, 1), |h_1|, |h_2| < M_0$, we have*

$$\mathbb{E} \langle \exp \lambda_0 N C_{1,2}^2 \rangle_0 \leq L, \quad (10)$$

where L is a constant independent of everything. In addition, the following system of equations has a unique solution $(a_0, b_0) \in [0, 1]^2$,

$$h_1^2 + \frac{\partial_x \xi(a_0, b_0)}{\gamma} = \frac{a_0}{(1 - a_0)^2}, \quad h_2^2 + \frac{\partial_y \xi(a_0, b_0)}{1 - \gamma} = \frac{b_0}{(1 - b_0)^2} \quad (11)$$

and (a_0, b_0) satisfies

$$|a_N - a_0|, |b_N - b_0| \leq \frac{L \log^2 N}{N}. \quad (12)$$

Proof. Note that $e^{\max(x,y)} \leq e^x + e^y$ for any x, y . It suffices to show that

$$\mathbb{E} \langle \exp \lambda N (R_{1,2}^1 - a_N)^2 \rangle_0, \quad \mathbb{E} \langle \exp \lambda N (R_{1,2}^2 - b_N)^2 \rangle_0 \leq L.$$

Since the configuration space in the bipartite model is $S^{N_1} \times S^{N_2}$, $R_{1,2}^1$ depends only on $u^1, u^2 \in S^{N_1}$ and $R_{1,2}^2$ depends only on $v^1, v^2 \in S^{N_2}$, we have that

$$\begin{aligned} \mathbb{E} \langle \exp \lambda N (R_{1,2}^1 - a_N)^2 \rangle_0 &= \mathbb{E} \langle \exp \lambda N (R_{1,2}^1 - a_N)^2 \rangle_1^-, \\ \mathbb{E} \langle \exp \lambda N (R_{1,2}^2 - b_N)^2 \rangle_0 &= \mathbb{E} \langle \exp \lambda N (R_{1,2}^2 - b_N)^2 \rangle_2^-, \end{aligned}$$

where $\langle \cdot \rangle_1^-$ and $\langle \cdot \rangle_2^-$ stand for the Gibbs expectations associated respectively to the Hamiltonians

$$\begin{aligned} &\sum_{i=1}^N u_i \left(y_i \sqrt{\gamma_N^{-1} \partial_x \xi(a_N, b_N)} + h_1 \right), \quad \forall u \in S^{N_1} \\ &\sum_{j=1}^N v_j \left(z_j \sqrt{(1 - \gamma_N)^{-1} \partial_y \xi(a_N, b_N)} + h_2 \right), \quad \forall v \in S^{N_2}. \end{aligned} \quad (13)$$

For Hamiltonians like (13), they have been studied thoughtfully in Panchenko [20]. In particular, one can find that exactly the same argument with only some minor modifications as pages 1045 – 1046 in [20] implies that there exists some $M_0 > 0$ and $\lambda_0 > 0$ such that whenever $\xi(1, 1), |h_1|, |h_2| < M_0$,

$$\begin{aligned} \mathbb{E} \langle \exp \lambda_0 N (R_{1,2}^1 - a_N)^2 \rangle_1^- &\leq L, \\ \mathbb{E} \langle \exp \lambda_0 N (R_{1,2}^2 - b_N)^2 \rangle_2^- &\leq L, \end{aligned}$$

which gives (10). In addition, the equation (1.5) and Lemma 7 in [20] says that there exist a'_N and b'_N satisfying

$$\begin{aligned} h_1^2 + \frac{\partial_x \xi(a'_N, b_N)}{\gamma_N} &= \frac{a'_N}{(1 - a'_N)^2}, \\ h_2^2 + \frac{\partial_y \xi(a_N, b'_N)}{1 - \gamma_N} &= \frac{b'_N}{(1 - b'_N)^2} \end{aligned}$$

such that

$$|a_N - a'_N|, |b_N - b'_N| \leq \frac{L \log^2 N}{N}. \quad (14)$$

These amount to say that

$$\begin{aligned} \left| h_1^2 + \frac{\partial_x \xi(a_N, b_N)}{\gamma_N} - \frac{a_N}{(1 - a_N)^2} \right| &\leq \frac{L \log^2 N}{N}, \\ \left| h_2^2 + \frac{\partial_y \xi(a_N, b_N)}{1 - \gamma_N} - \frac{b_N}{(1 - b_N)^2} \right| &\leq \frac{L \log^2 N}{N}. \end{aligned} \quad (15)$$

Define

$$\begin{aligned} F_1(a, b) &= (1-a)^2 \left(h_1^2 + \frac{\partial_x \xi(a, b)}{\gamma} \right), \\ F_2(a, b) &= (1-b)^2 \left(h_2^2 + \frac{\partial_y \xi(a, b)}{1-\gamma} \right), \\ F(a, b) &= (F_1(a, b), F_2(a, b)) \end{aligned}$$

for $(a, b) \in [0, 1] \times [0, 1]$. Since $h_1, h_2, \xi(1, 1)$ are all small, F maps into itself and a direct computation shows that $\sup_{a, b \in [0, 1]} \|DF(a, b)\| < 1$. So F forms a contraction mapping and there exists a unique $(a_0, b_0) \in [0, 1]^2$ such that $F(a_0, b_0) = (a_0, b_0)$. Note that here none of a_0 and b_0 could be equal to 1. To sum up, we obtain (11). On the other hand, since $G(a, b) := (F_1(a, b)/(1-a)^2, F_2(a, b)/(1-b)^2)$ for $(a, b) \in [0, 1]^2$ satisfies

$$DG(a, b) = \begin{bmatrix} \frac{\partial_{xx}\xi(a, b)}{\gamma} - \frac{1+a}{(1-a)^3} & \frac{\partial_{xy}\xi(a, b)}{\gamma} \\ \frac{\partial_{xy}\xi(a, b)}{1-\gamma} & \frac{\partial_{yy}\xi(a, b)}{1-\gamma} - \frac{1+b}{(1-b)^3} \end{bmatrix}$$

and $\xi(1, 1)$ is small, this ensures that DG is invertible in a small open neighborhood of (a_0, b_0) . This combining with (15) and the assumption that $|\gamma_N - \gamma| \leq K/N$ gives (12). \square

Recall the definition of $M = \max_{x, y \in [0, 1]} |\nabla^2 \xi(x, y)|$. Using this lemma, we take $\xi(1, 1) < M_0$ to be small enough such that $16M < \lambda_0$. Then for any $0 \leq t \leq 1$, $\lambda_0 - 8Mt \geq 16M - 8Mt \geq 8M$ and consequently, from (9),

$$\begin{aligned} \mathbb{E} \langle \exp 8MNC_{1,2}^2 \rangle_t &\leq \mathbb{E} \langle \exp(\lambda_0 - 8Mt) NC_{1,2}^2 \rangle_t \\ &\leq \mathbb{E} \langle \exp \lambda_0 NC_{1,2}^2 \rangle_0 \\ &\leq L. \end{aligned}$$

In other words, from $x^2 \leq 2e^{x^2}$ for all $x \in \mathbb{R}$, we obtain that for all $0 \leq t \leq 1$,

$$\mathbb{E} \langle C_{1,2}^2 \rangle_t \leq \frac{L}{N} \tag{16}$$

and therefore,

$$\left| \phi(1) - \phi(0) - \frac{1}{2} \Delta(1, 1) \right| \leq \frac{L}{N}. \tag{17}$$

Now note that $\phi(1) = N^{-1} \mathbb{E} \log Z_N$ and

$$\begin{aligned} \phi(0) &= \frac{1}{N} \mathbb{E} \log \int_{S^{N_1}} \exp \sum_{i=1}^{N_1} u_i \left(y_i \sqrt{\gamma_N^{-1} \partial_x \xi(a_N, b_N)} + h_1 \right) d\lambda_{N_1} \\ &\quad + \frac{1}{N} \mathbb{E} \log \int_{S^{N_2}} \exp \sum_{j=1}^{N_2} v_j \left(z_j \sqrt{(1-\gamma_N)^{-1} \partial_y \xi(a_N, b_N)} + h_2 \right) d\lambda_{N_2} \end{aligned}$$

Since $\lim_{N \rightarrow \infty} a_N = a_0$, $\lim_{N \rightarrow \infty} b_N = b_0$ and $\lim_{N \rightarrow \infty} \gamma_N = \gamma$, these guarantee that

$$\begin{aligned} \lim_{N \rightarrow \infty} \phi(0) &= \gamma \lim_{N \rightarrow \infty} \frac{1}{N_1} \mathbb{E} \log \int_{S^{N_1}} \exp \sum_{i=1}^{N_1} u_i \left(y_i \sqrt{\gamma^{-1} \partial_x \xi(a_0, b_0)} + h_1 \right) d\lambda_{N_1} \\ &\quad + (1 - \gamma) \lim_{N \rightarrow \infty} \frac{1}{N_2} \mathbb{E} \log \int_{S^{N_2}} \exp \sum_{j=1}^{N_2} v_j \left(z_j \sqrt{(1 - \gamma)^{-1} \partial_y \xi(a_0, b_0)} + h_2 \right) d\lambda_{N_2}. \end{aligned} \quad (18)$$

Here the two limits on the right-hand side are known as the simplest cases (the Hamiltonians are defined only by the external field) of the limiting free energy for the spherical SK model and are given respectively by the Crisanti-Sommers formula [22, Theorem 1.1 and Equation (1.11) therein],

$$\begin{aligned} &\frac{1}{2} \inf_{a \in [0,1]} \left\{ h_1^2(1 - a) + \frac{a}{1 - a} + \log(1 - a) + \frac{(1 - a) \partial_x \xi(a_0, b_0)}{\gamma} \right\}, \\ &\frac{1}{2} \inf_{b \in [0,1]} \left\{ h_2^2(1 - b) + \frac{b}{1 - b} + \log(1 - b) + \frac{(1 - b) \partial_y \xi(a_0, b_0)}{1 - \gamma} \right\}. \end{aligned} \quad (19)$$

Here the critical point equations in both cases are given by

$$\begin{aligned} h_1^2 + \frac{\partial_x \xi(a_0, b_0)}{\gamma} &= \frac{a}{(1 - a)^2}, \\ h_2^2 + \frac{\partial_y \xi(a_0, b_0)}{1 - \gamma} &= \frac{b}{(1 - b)^2}. \end{aligned}$$

Since (a_0, b_0) also satisfies (11), it follows $(a, b) = (a_0, b_0)$. This combining with (12), (17), (18) and (19) yields

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \log Z_N &= \frac{\gamma}{2} \left(h_1^2(1 - a_0) + \frac{a_0}{1 - a_0} + \log(1 - a_0) \right) + \frac{(1 - a_0) \partial_x \xi(a_0, b_0)}{2} \\ &\quad + \frac{(1 - \gamma)}{2} \left(h_2^2(1 - b_0) + \frac{b_0}{1 - b_0} + \log(1 - b_0) \right) + \frac{(1 - b_0) \partial_y \xi(a_0, b_0)}{2} \\ &\quad + \frac{1}{2} (\xi(1, 1) - \xi(a_0, b_0) - (1 - a_0) \partial_x \xi(a_0, b_0) - (1 - b_0) \partial_y \xi(a_0, b_0)) \\ &= P(a_0, b_0). \end{aligned}$$

Since

$$\nabla P(a, b) = \frac{1}{2} \left(-\gamma h_1^2 + \frac{\gamma a}{(1 - a)^2} - \partial_x \xi(a, b), -(1 - \gamma) h_2^2 + \frac{(1 - \gamma) b}{(1 - b)^2} - \partial_y \xi(a, b) \right),$$

it means that (a_0, b_0) is the only critical point of P . On the other hand, since

$$\nabla^2 P(a, b) = \frac{1}{2} \begin{pmatrix} \frac{\gamma(1+a)}{(1-a)^3} - \partial_{xx} \xi(a, b) & -\partial_{xy} \xi(a, b) \\ -\partial_{xy} \xi(a, b) & \frac{(1-\gamma)(1+b)}{(1-b)^3} - \partial_{yy} \xi(a, b) \end{pmatrix},$$

$a, b \in [0, 1)$ and $\xi(1, 1)$ is very small, we see that $\nabla^2 P(a, b)$ is positive semi-definite on $[0, 1)^2$. Therefore, (a_0, b_0) attains the infimum of P . This gives (3). As for (4), it can be obtained from (14) and (16).

3 Complexity of a coupled Hamiltonian

As the first step to prove Theorem 2, we will need to investigate the complexity of a coupled Hamiltonian using two independent pure p -spin spherical models. We believe the result here is also of interest by itself. Consider two independent mixed-spin spherical Hamiltonians $(\mathcal{H}_{N_1}^1(u) : u \in S^{N_1})$ and $(\mathcal{H}_{N_2}^2(v) : v \in S^{N_2})$ with mean zero and variance,

$$\mathbb{E}\mathcal{H}_{N_1}^1(u)\mathcal{H}_{N_1}^1(u') = N_1\xi^1(R^1(u, u')) \text{ and } \mathbb{E}\mathcal{H}_{N_2}^2(v)\mathcal{H}_{N_2}^2(v') = N_2\xi^2(R^2(v, v')), \quad (20)$$

where $\xi^1(x) := \sum_{p \geq 1} \beta_p^2 x^p$ and $\xi^2(y) := \sum_{q \geq 1} \beta_q'^2 y^q$. As before, we will assume that the sequences $(\beta_p)_{p \geq 1}$ and $(\beta_q')_{q \geq 1}$ decay sufficiently fast to zero so that $\mathcal{H}_{N_1}^1$ and $\mathcal{H}_{N_2}^2$ are smooth almost surely and we assume $\xi_1(1) = 1 = \xi_2(1)$. The coupled Hamiltonian we will be considering is defined as

$$\tilde{H}_N(u, v) = \mathcal{H}_{N_1}^1(u) + \mathcal{H}_{N_2}^2(v). \quad (21)$$

For this coupled Hamiltonian we can precisely derive the asymptotic complexity. We will use $\text{Crt}_{N,k}(t)$ for the variable defined in (5) with H_N replaced by \tilde{H}_N . Our main result in this direction is the following.

Theorem 3. *For any $k \in \mathbb{N} \cup \{0\}$, the following limit exists,*

$$\Upsilon_k(t) := \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \text{Crt}_{N,k}(t) < \infty.$$

Proof. For simplicity, we will assume $N_1 = \gamma N$ and $N_2 = (1 - \gamma)N$. All the arguments below hold under minor changes to (1). We fix an orthonormal basis $\Omega(u, v)$ of the tangent plane of $S^{N_1} \times S^{N_2}$ at (u, v) such that $\Omega(u, v) = (\Omega_1(u), \Omega_2(v)) = (\Omega_1, \Omega_2)$, where $\Omega_i = \{E_i^1, \dots, E_i^{N_i-1}\}$ is an orthonormal basis of S^{N_i} at u (or v) respectively. We choose $\Omega(u, v)$ continuously on (u, v) . In view of (21), one can write

$$\nabla \tilde{H}_N(u, v) = (\nabla \mathcal{H}_{N_1}^1(u), \nabla \mathcal{H}_{N_2}^2(v)) \quad (22)$$

and

$$\nabla^2 \tilde{H}_N(u, v) = \begin{pmatrix} \nabla^2 \mathcal{H}_{N_1}^1(u) & 0 \\ 0 & \nabla^2 \mathcal{H}_{N_2}^2(v) \end{pmatrix}. \quad (23)$$

We now write the event $\{i(\nabla^2 \tilde{H}_N(u, v)) = k\}$ as the following disjoint union

$$\left\{i(\nabla^2 \tilde{H}_N(u, v)) = k\right\} = \bigcup_{l=0}^k \left\{i(\nabla^2 \mathcal{H}_{N_1}^1(u)) = l, i(\nabla^2 \mathcal{H}_{N_2}^2(v)) = k - l\right\} = \bigcup_{l=0}^k (E_l^1 \cap E_{k-l}^2),$$

which leads to

$$\mathbb{E} \text{Crt}_{N,k}(t) = \sum_{l=0}^k \mathbb{E} \text{Crt}_{N,k}(t) \mathbf{1}_{E_l^1} \mathbf{1}_{E_{k-l}^2}.$$

Conditioning on the value of $\mathcal{H}_{N_1}^1$, each term of the sum above can be written as

$$\begin{aligned} & \frac{\sqrt{N_1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbb{E}[\#\{(u, v) : \nabla \mathcal{H}_{N_1}^1(u) = 0, \nabla \mathcal{H}_{N_2}^2(v) = 0, \\ & \quad \mathcal{H}_{N_2}^2(v) \leq Nt - N_1x, E_l^1, E_{k-l}^2\} | \mathcal{H}_{N_1}^1(u) = N_1x] e^{-\frac{N_1x^2}{2}} dx \\ &= \frac{\sqrt{N_1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbb{E}[\text{Crt}_{N_2, k-l}^{\mathcal{H}_{N_2}^2}((1-\gamma)^{-1}(t-\gamma x))] \\ & \quad \cdot \mathbb{E}[\#\{u : \nabla \mathcal{H}_{N_1}^1(u) = 0, E_l^1\} | \mathcal{H}_{N_1}^1(u) = N_1x] e^{-\frac{N_1x^2}{2}} dx \end{aligned}$$

Now from [5, Theorem 1.1] (see also [6, Theorem 2.5] for the pure case), there exist functions Θ_l and Λ_k from \mathbb{R} to $\mathbb{R} \cup \{-\infty\}$, both bounded above and smooth when finite, such that

$$\begin{aligned} \mathbb{E}[\text{Crt}_{N_2, k-l}^{\mathcal{H}_{N_2}^2}((1-\gamma)^{-1}(t-\gamma x))] &\sim \exp(N_2 \Theta_{k-l}((1-\gamma)^{-1}(t-\gamma x)) \\ \mathbb{E}[\#\{u : \nabla \mathcal{H}_{N_1}^1(u) = 0, E_l^1\} | \mathcal{H}_{N_1}^1(u) = N_1x] &\sim \exp(N_1 \Lambda_l(x)) \end{aligned}$$

where $a_n(x) \sim b_n(x)$ means $\lim n^{-1} \log(a_n(x)/b_n(x)) = 0$ uniformly in x . This allows us to replace the above limits in (21) to obtain after an application of Laplace's Method the expression

$$\Upsilon_k(t) = \sup_{0 \leq l \leq k} \sup_{x \in \mathbb{R}} \left[(1-\gamma) \Theta_{k-l} \left(\frac{t-\gamma x}{1-\gamma} \right) + \gamma \left(\Lambda_l(x) - \frac{x^2}{2} \right) \right]. \quad (24)$$

□

In the case of the pure p, q Hamiltonian, one has for instance the expression (see [6, Equation (2.16)])

$$\begin{aligned} \Upsilon_0(t) = \sup_{x \in \mathbb{R}} \left\{ (1-\gamma) \left(\frac{1}{2} \log(q-1) - \frac{q-2}{4(q-1)} \frac{(t-\gamma x)^2}{(1-\gamma)^2} - I\left(\frac{t-\gamma x}{1-\gamma}\right) \right) \right. \\ \left. + \gamma \left(\frac{1}{2} \log(p-1) - \frac{p-2}{4(p-1)} x^2 - I(x) \right) \right\} \end{aligned}$$

where the function I is defined in (42).

4 Proof of Theorem 2

Recall $\mathcal{H}_{N_1}^1$ and $\mathcal{H}_{N_2}^2$ from (20) and $\xi(x, y)$ from (2). From now on, we will choose

$$\beta_p^2 = \sum_{q \geq 1} \beta_{p,q}^2 \text{ and } \beta_q'^2 = \sum_{p \geq 1} \beta_{p,q}^2 \quad (25)$$

for all $p, q \geq 1$. To lighten notations, we will simply denote H_N as H and \tilde{H}_N as \tilde{H} . The hypothesis on H and \tilde{H} allows us to apply Rice's formula in the form of Lemma 3.1 of [6],

which says

$$\begin{aligned}\mathbb{E} \text{Crt}_{N,k}^H(t) &= V_{N_1} V_{N_2} \int_{S^{N_1} \times S^{N_2}} \zeta(u, v) d\lambda_{N_1} \times \lambda_{N_2}(u, v), \\ \mathbb{E} \text{Crt}_{N,k}^{\tilde{H}}(t) &= V_{N_1} V_{N_2} \int_{S^{N_1} \times S^{N_2}} \tilde{\zeta}(u, v) d\lambda_{N_1} \times \lambda_{N_2}(u, v),\end{aligned}\tag{26}$$

where

$$V_N = \frac{2(\pi N)^{N/2}}{\Gamma(N/2)}$$

and

$$\begin{aligned}\zeta(u, v) &:= \mathbb{E}[|\det \nabla^2 H(u, v)| \mathbf{1}\{H(u, v) \leq Nt, i(\nabla^2 H(u, v)) = k\} \mid \nabla H(u, v) = 0] \phi_{(u,v)}(0), \\ \tilde{\zeta}(u, v) &:= \mathbb{E}[|\det \nabla^2 \tilde{H}(u, v)| \mathbf{1}\{\tilde{H}(u, v) \leq Nt, i(\nabla^2 \tilde{H}(u, v)) = k\} \mid \nabla \tilde{H}(u, v) = 0] \tilde{\phi}_{(u,v)}(0).\end{aligned}\tag{27}$$

Here $\phi_{(u,v)}$ and $\tilde{\phi}_{(u,v)}$ are the joint densities of the gradient vectors of H and \tilde{H} at (u, v) , respectively. Due to rotation symmetry of the law of the Hamiltonian H_N , we have

Lemma 2. *For any $u_1, u_2 \in S^{N_1}$ and $v_1, v_2 \in S^{N_2}$, we have $\zeta(u_1, v_1) = \zeta(u_2, v_2)$ and $\tilde{\zeta}(u_1, v_1) = \tilde{\zeta}(u_2, v_2)$.*

Proof. Let θ_1 and θ_2 be orthogonal transformations of S^{N_1} and S^{N_2} that send u_1 to u_2 and v_1 to v_2 , respectively. Recall the continuous orthonormal basis Ω from the proof of Theorem 3. Write θE for the image of Ω under the rotation (θ_1, θ_2) . Then $\hat{H}(u, v) := H(\theta_1 u, \theta_2 v)$ has the same law as H since they share the same mean and covariance function. Although it is not necessarily true that $\theta_1 E_i^1(\theta_1 u, \theta_2 v) = E_i^1(\theta_1 u, \theta_2 v)$ or $\theta_2 E_i^2(\theta_1 u, \theta_2 v) = E_i^2(\theta_1 u, \theta_2 v)$, the value of $|\det \nabla^2 \hat{H}(u, v)|$ is the same regardless of the choice of the basis at the tangent plane. Furthermore, the law of $\nabla H(u, v)$ is the same as the law of $\nabla \hat{H}(u, v)$ (written in the θE coordinates) and being a critical point is independent of the choice of the basis at the tangent plane. These facts establish that $\zeta(u_1, v_1) = \zeta(u_2, v_2)$ and same arguments hold true for $\tilde{\zeta}$. □

By this lemma, it suffices to understand the functions ζ and $\tilde{\zeta}$ at the "double north pole"

$$\mathbf{n} := (0, \dots, 0, \sqrt{N_1}, 0, \dots, 0, \sqrt{N_2}).$$

Consequently, since $\lambda_{N_1} \times \lambda_{N_2}$ is a probability measure, we have

$$\begin{aligned}\mathbb{E} \text{Crt}_{N,k}^H(t) &= V_{N_1} V_{N_2} \zeta(\mathbf{n}), \\ \mathbb{E} \text{Crt}_{N,k}^{\tilde{H}}(t) &= V_{N_1} V_{N_2} \tilde{\zeta}(\mathbf{n}).\end{aligned}\tag{28}$$

To proceed we will need one lemma. It establishes the covariance structure of the vector $(H(\mathbf{n}), \nabla H(\mathbf{n}), \nabla^2 H(\mathbf{n}))$. We use the notation given in (22) and (23). Since the difference between H and \tilde{H} will be mainly present on the non-diagonal blocks of their Hessians, we also include the covariance structure for $(\tilde{H}(\mathbf{n}), \nabla \tilde{H}(\mathbf{n}), \nabla^2 \tilde{H}(\mathbf{n}))$ to emphasize to the reader the

difference between the coupled Hamiltonian \tilde{H} and the Hamiltonian of the bipartite model. Recall the function $\xi(x, y)$ from (4). Set

$$\xi'_1 = \partial_x \xi(1, 1), \xi'_2 = \partial_y \xi(1, 1), \xi''_1 = \partial_x^2 \xi(1, 1), \xi''_2 = \partial_y^2 \xi(1, 1)$$

and for $i = 1, 2$,

$$\alpha_i = (\xi''_i + \xi'_i - \xi_i'^2)^{1/2}. \quad (29)$$

Here the quantity in the square root of (29) is nonnegative as can be easily verified using $\xi(1, 1) = 1$ and the Cauchy-Schwarz inequality. So α_i is well-defined. Set $I_1 = \{1, 2, \dots, N_1 - 1\}$ and $I_2 = \{1, 2, \dots, N_2 - 1\}$.

Lemma 3. *Suppose that $X = H$ or \tilde{H} . Write*

$$\nabla X(\mathbf{n}) = (\nabla X^1(\mathbf{n}), X^2(\mathbf{n}))$$

and

$$\nabla^2 X(\mathbf{n}) = \begin{pmatrix} \nabla^2 X^{11}(\mathbf{n}) & \nabla^2 X^{12}(\mathbf{n}) \\ \nabla^2 X^{21}(\mathbf{n}) & \nabla^2 X^{22}(\mathbf{n}) \end{pmatrix}$$

as in (22) and (23). The following statements hold.

- 1 The gradient $\nabla X(\mathbf{n})$ is independent of the vector $(X(\mathbf{n}), \nabla^2 X(\mathbf{n}))$. Furthermore, $\nabla X(\mathbf{n})$ is a vector of $N - 2$ independent Gaussian random variables with variance

$$\mathbb{E}[\nabla X_i^1(\mathbf{n}) \nabla X_i^1(\mathbf{n})] = \frac{\xi'_1 N}{N_1}, \quad \mathbb{E}[\nabla X_j^2(\mathbf{n}) \nabla X_j^2(\mathbf{n})] = \frac{\xi'_2 N}{N_2}, \quad \forall (i, j) \in I_1 \times I_2. \quad (30)$$

- 2 The only correlated entries of the matrix $\nabla^2 X^{11}(\mathbf{n})$ are in the diagonal. Precisely, for $i, i' \in I_1$ with $i \neq i'$,

$$\begin{aligned} \text{Var}[\nabla^2 X_{ii}^{11}(\mathbf{n})] &= \frac{N}{N_1^2} (\xi'_1 + 3\xi''_1), \\ \text{Var}[\nabla^2 X_{ii'}^{11}(\mathbf{n})] &= \frac{N}{N_1^2} \xi''_1, \\ \mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n}) \nabla^2 X_{i'i'}^{11}(\mathbf{n})] &= \frac{N}{N_1^2} (\xi'_1 + \xi''_1). \end{aligned}$$

- 3 The covariance structure of the matrix $\nabla^2 X^{22}(\mathbf{n})$ is as in the item 2 above replacing ξ'_1, ξ''_1 by ξ'_2, ξ''_2 and N_1 by N_2 .

- 4 The diagonal entries of $\nabla^2 X^{11}(\mathbf{n})$, $\nabla^2 X^{22}(\mathbf{n})$ are correlated as

$$\mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n}) \nabla^2 X_{jj}^{22}(\mathbf{n})] = \frac{N}{N_1 N_2} \xi'_1 \xi'_2, \quad \forall (i, j) \in I_1 \times I_2.$$

All other correlations between entries of $\nabla^2 X^{11}(\mathbf{n})$ and $\nabla^2 X^{22}(\mathbf{n})$ are zero.

- 5 The off-diagonal block matrices $\nabla^2 X^{12}(\mathbf{n})$ and $\nabla^2 X^{21}(\mathbf{n})$ are independent of both $\nabla^2 X^{11}(\mathbf{n})$ and $\nabla^2 X^{22}(\mathbf{n})$. All their entries are independent and with variance given by: $\forall(i, j) \in I_1 \times I_2$,

$$\text{Var}[\nabla^2 X_{ij}^{12}(\mathbf{n})] = \text{Var}[\nabla^2 X_{ji}^{21}(\mathbf{n})] = \begin{cases} \frac{N}{N_1 N_2} \xi'_1 \xi'_2, & \text{if } X = H, \\ 0, & \text{if } X = \tilde{H}. \end{cases}$$

- 6 The off-diagonal entries of $\nabla^2 X(\mathbf{n})$ are independent of X . The diagonal terms satisfy

$$\mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n})X(\mathbf{n})] = -\frac{N}{N_1} \xi'_1, \quad \mathbb{E}[\nabla^2 X_{jj}^{22}(\mathbf{n})X(\mathbf{n})] = -\frac{N}{N_2} \xi'_2, \quad \forall(i, j) \in I_1 \times I_2.$$

- 7 Conditioning on $X(\mathbf{n}) = x$, the off-diagonal entries of the random matrix $\nabla^2 X(\mathbf{n})$ are independent centered Gaussian random variables. Furthermore, the diagonal entries have mean

$$\mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n})] = -\frac{\xi'_1 x}{N_1}, \quad \mathbb{E}[\nabla^2 X_{jj}^{22}(\mathbf{n})] = -\frac{\xi'_2 x}{N_2}, \quad \forall(i, j) \in I_1 \times I_2$$

and for $i, i' \in I_1$ and $j, j' \in I_2$,

$$\begin{aligned} \mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n})\nabla^2 X_{i'i'}^{11}(\mathbf{n})] &= \frac{N}{N_1^2}(\xi'_1 + (1 + 2\delta_{ii'})\xi''_1 - \xi'^2_1), \quad \text{Var}[\nabla^2 X_{ii}^{11}(\mathbf{n})] = \frac{N}{N_1^2}\xi''_1, \\ \mathbb{E}[\nabla^2 X_{jj}^{22}(\mathbf{n})\nabla^2 X_{j'j'}^{22}(\mathbf{n})] &= \frac{N}{N_2^2}(\xi'_2 + (1 + 2\delta_{jj'})\xi''_2 - \xi'^2_2), \quad \text{Var}[\nabla^2 X_{jj}^{22}(\mathbf{n})] = \frac{N}{N_2^2}\xi''_2, \\ \mathbb{E}[\nabla^2 X_{ii}^{11}(\mathbf{n})\nabla^2 X_{jj}^{22}(\mathbf{n})] &= \frac{N}{N_1 N_2} \xi'_1 \xi'_2, \end{aligned}$$

$$\text{Var}[\nabla^2 X_{ii'}^{12}(\mathbf{n})] = \text{Var}[\nabla^2 X_{jj'}^{21}(\mathbf{n})] = \begin{cases} \frac{N}{N_1 N_2} \xi'_1 \xi'_2 & \text{if } X = H, \\ 0, & \text{if } X = \tilde{H}. \end{cases}$$

Proof. The proof basically follows from standard computations involving derivatives of smooth Gaussian fields as in [7, Lemma 8.5]. We will only provide the proof for $X = H$. As for the case $X = \tilde{H}$, it can be treated in a similar manner by noting that (25) and the functions $\xi^1(x) = \sum_{p \geq 1} \beta_p^2 x^p$ and $\xi^2(x) = \sum_{q \geq 1} \beta_q'^2 x^q$ appeared in the definition of \tilde{H} satisfy

$$\xi^i(1) = 1, \quad \xi^{i'}(1) = \xi'_i, \quad \xi^{i''}(1) = \xi''_i$$

for $i = 1, 2$. First, we define the function $\Psi_k : S^k \rightarrow \mathbb{R}^{k-1}$ by $\Psi_k(x_1, \dots, x_k) = (x_1, \dots, x_{k-1})$. The function $\Psi : S^{N_1} \times S^{N_2} \rightarrow \mathbb{R}^{N-2}$ given by

$$\Psi(u, v) = (\Psi_{N_1} u, \Psi_{N_2} v)$$

is a chart in some neighborhood U of \mathbf{n} . We set

$$\tilde{H} = H \circ \Psi^{-1},$$

which is a Gaussian process on $\Psi(U)$ with covariance

$$\begin{aligned}
& C((u, v), (w, z)) \\
&= \text{Cov}(\bar{H}(u, v), \bar{H}(w, z)) \\
&= \sum_{p, q \geq 1} \frac{\beta_{p, q} N}{N_1^p N_2^q} \left\{ \sum_{i=1}^{N_1-1} u_i w_i + \sqrt{(N_1 - \sum_{i=1}^{N_1-1} u_i^2)(N_1 - \sum_{i=1}^{N_1-1} w_i^2)} \right\}^p \\
&\quad \cdot \left\{ \sum_{i=1}^{N_2-1} v_i z_i + \sqrt{(N_2 - \sum_{i=1}^{N_2-1} v_i^2)(N_2 - \sum_{i=1}^{N_2-1} z_i^2)} \right\}^q.
\end{aligned} \tag{31}$$

As in the proof of Theorem 3, we chose the orthogonal basis of the tangent spaces to coincide with the vector field $E_i, i = 1 \dots, N-2$. Then the covariant Hessian $\nabla^2 H(\mathbf{n})$ coincides with the Euclidean Hessian of $\bar{H}(0, 0)$, by noting that the Christoffel symbols $\Gamma_{kl}^i(\mathbf{n}) \equiv 0$. The covariances of $\bar{H}, \nabla \bar{H}, \nabla^2 \bar{H}$, are computed using a well-known formula [1, Formula (5.5.5)]

$$\mathbb{E} \left\{ \frac{\partial^{\alpha+\beta} \bar{H}(u, v)}{\partial^\alpha u \partial^\beta v} \frac{\partial^{\rho+\zeta} \bar{H}(u, v)}{\partial^\rho u \partial^\zeta v} \right\} = \frac{\partial^{\alpha+\beta+\rho+\zeta} C((u, v), (w, z))}{\partial^\alpha u \partial^\beta v \partial^\rho w \partial^\zeta z},$$

that relates the covariance of the derivatives of H to the derivatives of the covariance function C . Since the derivatives of a centered Gaussian field have centered Gaussian distribution, the covariance matrix determines uniquely the joint distribution of $(H, \nabla H, \nabla^2 H)$ and items 1 – 6 follow from [1, Formula (5.5.4)] applied to (31). For instance, item 1 follows from differentiating C with respect to u_i, w_j for all possible values of i and j and evaluating at $u = v = w = z = 0$. Item 7 follows from the rules on Gaussian distributions transform under conditioning (see, e.g., [1, Pages 10-11]).

□

Write I_m for the $m \times m$ identity matrix. Item 7 of the lemma above says that under the law $\mathbb{P}[\cdot | H(\mathbf{n}) = x]$, the matrix $\nabla^2 H(\mathbf{n})$ has the same distribution as the random matrix

$$\begin{pmatrix} G_1 & G \\ G^t & G_2 \end{pmatrix} - \begin{pmatrix} \frac{\xi'_1 x}{N_1} I_{N_1} & 0 \\ 0 & \frac{\xi'_2 x}{N_2} I_{N_2} \end{pmatrix}. \tag{32}$$

Here, G is a $(N_1 - 1) \times (N_2 - 1)$ matrix with independent Gaussian entries with mean zero and variance $N \xi'_1 \xi'_2 / (N_1 N_2)$ and for $i = 1, 2$,

$$G_i \stackrel{d}{=} \left(\frac{N(N_i - 1)}{N_i^2} 2\xi''_i \right)^{1/2} M^{N_i-1} + \frac{\sqrt{N}}{N_i} \alpha_i Z_i I_{N_i}, \tag{33}$$

where α_i is defined by (29), Z_i 's are independent standard Gaussian random variables and M^{N_i-1} is a $(N_i - 1) \times (N_i - 1)$ GOE matrix (see (39) below). To proceed, under the law $\mathbb{P}[\cdot | H(\mathbf{n}) = x]$, we consider the decomposition $H = H_1 + H_2$ for $a \in [0, 1]$, where H_1 and H_2 are independent random matrices defined as

$$\begin{aligned}
H_1 &= H_1(a) = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} - a \begin{pmatrix} \frac{\xi'_1 x}{N_1} I_{N_1} & 0 \\ 0 & \frac{\xi'_2 x}{N_2} I_{N_2} \end{pmatrix}, \\
H_2 &= H_2(a) = \begin{pmatrix} 0 & G \\ G^t & 0 \end{pmatrix} - (1-a) \begin{pmatrix} \frac{\xi'_1 x}{N_1} I_{N_1} & 0 \\ 0 & \frac{\xi'_2 x}{N_2} I_{N_2} \end{pmatrix}.
\end{aligned} \tag{34}$$

Remark 2. This is the part where the computation for H and \tilde{H} goes in different ways. The reader familiar with the computations performed in [5] and [6] can foresee that modulo some small technical points the computation for \tilde{H} follows directly the path taken for the mixed p -spin model. The reason is that for \tilde{H} the matrix G above is identically zero and the Hessian becomes a diagonal block matrix where each block is exactly the Hessian of a mixed p -spin model (see the proof of Theorem 3). Therefore, one does not need the matrix H_2 in (34). However, the Hamiltonian H of the bipartite model does not share this property and one needs to deal with the off diagonal block matrix. As far as we know, exact formulas are not available and one cannot perform exact computations. The alternative taken here is to try to bound the (determinant of) Hessian H by its block terms using the decomposition (34), where the constant a plays the weight assigned to each part of the sum. Optimizing over the choice of a will lead to Theorem 2.

Recall the function ζ defined in (27) and consider the case $k = 0$. Note that $H(\mathbf{n})$ is centered Gaussian with variance $N\xi(1,1) = N$. We first condition on the value of H and use item 1 of Lemma 3 to write

$$\begin{aligned}\zeta(\mathbf{n}) &= \frac{\phi_{\mathbf{n}}(0)\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^t \mathbb{E}[|\det \nabla^2 H(\mathbf{n})| \mathbf{1}_{\{i(\nabla^2 H(\mathbf{n}))=0\}} \mid H(\mathbf{n}) = Nx] e^{-\frac{Nx^2}{2}} dx, \\ \tilde{\zeta}(\mathbf{n}) &= \frac{\tilde{\phi}_{\mathbf{n}}(0)\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^t \mathbb{E}[|\det \nabla^2 \tilde{H}(\mathbf{n})| \mathbf{1}_{\{i(\nabla^2 \tilde{H}(\mathbf{n}))=0\}} \mid \tilde{H}(\mathbf{n}) = Nx] e^{-\frac{Nx^2}{2}} dx,\end{aligned}\tag{35}$$

where recalling that $\phi_{\mathbf{n}}$ and $\tilde{\phi}_{\mathbf{n}}$ are the densities for H and \tilde{H} respectively, item 1 in Lemma 3 implies that they are both centered Gaussian with covariance structures (30) and thus,

$$\phi_{\mathbf{n}}(0) = \tilde{\phi}_{\mathbf{n}}(0) = \left((2\pi N)^{N-2} \left(\frac{\xi'_1}{N_1} \right)^{N_1-1} \left(\frac{\xi'_2}{N_2} \right)^{N_2-1} \right)^{-1/2}.\tag{36}$$

The control of $\mathbb{E} \text{Crt}_{N,0}^{H_N}(t)$ from above and below proceeds is discussed in the following two subsections.

4.1 Upper bound

Take the decomposition (34) with $a = 1$. Recall that Fischer's inequality ([14, Fact 8.11.26]) says that if

$$\mathcal{A} = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix}$$

is positive definite, then $\det \mathcal{A} \leq \det A \det C$. From this, we obtain

$$|\det \nabla^2 H(\mathbf{n})| \mathbf{1}_{\{i(\nabla^2 H(\mathbf{n}))=0\}} \leq \prod_{i=1}^2 |\det \nabla^2 H^{ii}(\mathbf{n})| \mathbf{1}_{\{i(\nabla^2 H^{ii}(\mathbf{n}))=0\}}.$$

In view of (35), applying item 7 in Lemma 3 for both $H(\mathbf{n})$ and $\tilde{H}(\mathbf{n})$ and using (36) lead to

$$\begin{aligned}\zeta(\mathbf{n}) &\leq \frac{\phi_{\mathbf{n}}(0)\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^t \prod_{i=1}^2 \mathbb{E} \left[\det \left(G_i - \frac{\xi'_i x}{N_i} I_{N_i} \right) \mathbf{1}_{\left\{ G_i - \frac{\xi'_i x}{N_i} I_{N_i} \geq 0 \right\}} \right] e^{-\frac{x^2}{2N}} dx \\ &= \frac{\phi_{\mathbf{n}}(0)\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^t \mathbb{E}[\det \tilde{H}_N(\mathbf{n}) \mathbf{1}_{\{\nabla^2 \tilde{H}_N(\mathbf{n}) \geq 0\}} | \tilde{H}_N(\mathbf{n}) = Nx] e^{-\frac{x^2}{2N}} dx \\ &= \tilde{\zeta}(\mathbf{n}).\end{aligned}$$

From the definition (26) and (27) of the complexity functions for $H(\mathbf{n})$ and $\tilde{H}(\mathbf{n})$, we get $\mathbb{E} \text{Crt}_{N,0}^H(t) \leq \mathbb{E} \text{Crt}_{N,0}^{\tilde{H}}(t)$ for all N and all t . From Theorem 3, we obtain the upper bound with $K(t) = \Upsilon_0(t)$. The fact that $\lim_{t \rightarrow -\infty} K(t) = -\infty$ follows from (24) by noting that $\lim_{t \rightarrow -\infty} \Theta_0(t) = -\infty$ and Λ_0 is bounded above (see [5, Proposition 1]).

4.2 Lower bound

Recall the decomposition H_1 and H_2 from (34) by replacing x by Nx . From item 7 in Lemma 3, under $\mathbb{P}[\cdot | H(\mathbf{n}) = Nx]$, $|\det \nabla^2 H| \mathbf{1}_{\{\nabla^2 H(\mathbf{n})=0\}}$ has the same law as $|\det(H_1 + H_2)| \mathbf{1}_{\{\nabla^2 H(\mathbf{n})=0\}}$. On the other hand,

$$\begin{aligned}|\det(H_1 + H_2)| \mathbf{1}_{\{\nabla^2 H(\mathbf{n})=0\}} &\geq |\det(H_1 + H_2)| \mathbf{1}_{\{H_1 \geq 0\}} \mathbf{1}_{\{H_2 \geq 0\}} \\ &\geq (\det H_1 + \det H_2) \mathbf{1}_{\{H_1 \geq 0\}} \mathbf{1}_{\{H_2 \geq 0\}},\end{aligned}$$

where the second inequality used the Minkowski determinant inequality [14, Corollary 8.4.15]. This implies that

$$\begin{aligned}&\mathbb{E} \left[|\det \nabla^2 H(\mathbf{n})| \mathbf{1}_{\{\nabla^2 H(\mathbf{n})=0\}} | H(\mathbf{n}) = Nx \right] \\ &\geq \sup_{a \in [0,1]} \mathbb{E} \left[\det H_1 \mathbf{1}_{\{H_1 \geq 0\}} \right] \mathbb{P}(H_2 \geq 0) \\ &= \sup_{a \in [0,1]} \left\{ \mathbb{P}(H_2 \geq 0) \prod_{i=1}^2 \mathbb{E} \left[\det \left(G_i - a \frac{N \xi'_i x}{N_i} I_{N_i} \right) \mathbf{1}_{\left\{ G_i - a \frac{N \xi'_i x}{N_i} I_{N_i} \geq 0 \right\}} \right] \right\}.\end{aligned}\tag{37}$$

Lemma 4. Let $\gamma_* = \max\{\gamma(1-\gamma)^{-1}, \gamma^{-1}(1-\gamma)\}$. If

$$(1-a) \min \left\{ \frac{\xi'_1 x}{\gamma}, \frac{\xi'_2 x}{(1-\gamma)} \right\} < -(1 + \sqrt{\gamma_*}),\tag{38}$$

then $\lim_{N \rightarrow \infty} \mathbb{P}(H_2 \geq 0) = 1$.

Proof. Denote by $\lambda_m(A)$ and $\lambda_M(A)$ the minimum and maximum eigenvalues for any symmetry matrix A . Note that H_2 is positive definite if and only if $\lambda_m(H_2) > 0$. Writing $\lambda_m(H_2) = \min_{v \neq 0} \langle H_2 v, v \rangle / \langle v, v \rangle$, this expression implies that $\lambda_m(H_2) > 0$ if

$$\lambda_m(\mathcal{G}) > c_N := (1-a) \min \left\{ \frac{N \xi'_1 x}{N_1}, \frac{N \xi'_2 x}{N_2} \right\}$$

for

$$\mathcal{G} := \begin{pmatrix} 0 & G \\ G^t & 0 \end{pmatrix}.$$

Let

$$c = \lim_{N \rightarrow \infty} c_N = (1 - a) \min \left\{ \frac{\xi'_1 x}{\gamma}, \frac{\xi'_2 x}{(1 - \gamma)} \right\}.$$

It is a well-known fact (see [3, Exercise 2.1.18]) that the largest eigenvalue of GG^t converges almost surely to the edge of the Marchenko-Pastur law, which in our normalization is given by $(1 + \sqrt{\gamma^*})^2$. Thus, with probability one, $\limsup_{N \rightarrow \infty} \lambda_M(GG^t) \leq (1 + \sqrt{\gamma^*})^2 + \varepsilon$ for any $\varepsilon > 0$. Using this and the fact that λ^2 is an eigenvalue of GG^t for any eigenvalue λ of \mathcal{G} , we obtain that if $c < -(1 + \sqrt{\gamma^*})$, then

$$c^2 + \varepsilon > (1 + \sqrt{\gamma^*})^2 + \varepsilon \geq \limsup_{N \rightarrow \infty} \lambda_M(GG^t) \geq \limsup_{N \rightarrow \infty} \lambda_m(\mathcal{G})^2$$

and letting $\varepsilon \downarrow 0$ gives $c^2 > \limsup_{N \rightarrow \infty} \lambda_m(\mathcal{G})^2$. Since G is formed by i.i.d. standard Gaussians, $\lambda_m(\mathcal{G})$ is almost surely negative. These amount to say that if $c < -(1 + \sqrt{\gamma^*})$, $c < \liminf_{N \rightarrow \infty} \lambda_m(\mathcal{G})$ with probability one. Using Fatou's lemma, we are done since

$$\begin{aligned} 1 &= \mathbb{P} \left(c < \liminf_{N \rightarrow \infty} \lambda_m(\mathcal{G}) \right) \\ &\leq \int \liminf_{N \rightarrow \infty} 1 \{ \lambda_m(\mathcal{G}) > c_N \} d\mathbb{P} \\ &\leq \liminf_{N \rightarrow \infty} \int 1 \{ \lambda_m(\mathcal{G}) > c_N \} d\mathbb{P} \\ &\leq \liminf_{N \rightarrow \infty} \mathbb{P}(H_2 \geq 0). \end{aligned}$$

□

This lemma implies that if we choose x and a such that (38) holds, then for N large enough $\mathbb{P}(H_2 \geq 0)$ is bounded below by a constant $C > 0$. Now we need to investigate the term

$$\prod_{i=1}^2 \mathbb{E} \left[\det \left(G_i - a \frac{N \xi'_i x}{N_i} I_{N_i} \right) \mathbf{1}_{\left\{ G_i - a \frac{N \xi'_i x}{N_i} I_{N_i} \geq 0 \right\}} \right]$$

for which we will need Lemma 3.3 [6]. Its precise statement is given below. Recall that a real symmetry matrix M^N of size $N \times N$ is a Gaussian Orthogonal Ensemble (GOE) if the entries $(M_{ij}^N, i \leq j)$ are independent Gaussian random variables with mean zero and variance

$$\mathbb{E}(M_{ij}^N)^2 = \frac{1 + \delta_{ij}}{2N}. \quad (39)$$

Denote by \mathbb{E}_G^N the expectation under the GOE of size $N \times N$.

Lemma 5. [6, Lemma 3.3] *Let M^{N-1} be a $(N-1) \times (N-1)$ GOE matrix, $\lambda_k^{(N-1)}$ be the k -th smallest eigenvalue of M^{N-1} and Y be an independent Gaussian random variable with*

mean m and variance t^2 . Then, for any Borel set $B \subset \mathbb{R}$,

$$\begin{aligned} & \mathbb{E} \left[\left| \det(M^{N-1} - YI) \right| \mathbf{1} \{ i(M^{N-1} - YI) = k, Y \in B \} \right] \\ &= \frac{\Gamma(\frac{N}{2})(N-1)^{-\frac{N}{2}}}{\sqrt{\pi t^2}} \mathbb{E}_G^N \left[\exp \left\{ \frac{N(\lambda_k^N)^2}{2} - \frac{\left(\left(\frac{N}{N-1} \right)^{\frac{1}{2}} \lambda_k^N - m \right)^2}{2t^2} \right\} \mathbf{1} \left\{ \lambda_k^N \in \left(\frac{N-1}{N} \right)^{\frac{1}{2}} B \right\} \right]. \end{aligned} \quad (40)$$

Assume first that $\alpha_i \neq 0$ for $i = 1, 2$. In the notation of Lemma 5, we take $B = \mathbb{R}$, $N = N_i$ and Y to be a Gaussian random variable of mean $(N(2\xi_i''(N_i - 1))^{-1})^{1/2} a \xi_i' x$ and variance $\alpha_i^2 (2\xi_i''(N_i - 1))^{-1}$. Recalling that $N_1/N = \gamma$, (40) together with (33) gives us the following lower bound for the right side of (37),

$$\frac{1}{2\pi} \prod_{i=1}^2 \frac{\Gamma(\frac{N_i}{2})(2\xi_i'')^{\frac{N_i+1}{2}}}{\alpha_i} \left(\frac{N}{N_i^2} \right)^{\frac{N_i-1}{2}} \mathbb{E}_G^{N_i} \left[\exp \left\{ \frac{N_i}{2} \left((\lambda_0^{N_i})^2 - \frac{2\xi_i''}{\alpha_i^2} \left(\lambda_0^{N_i} - \frac{a \xi_i' x}{\gamma_i (2\xi_i'')^{\frac{1}{2}}} \right)^2 \right) \right\} \right],$$

where $\gamma_1 := \gamma$, $\gamma_2 := 1 - \gamma$ and $\lambda_0^{N_1}$ and $\lambda_0^{N_2}$ are the smallest eigenvalues of two independent GOE of sizes N_1 and N_2 , respectively. Now combining the equation above with (28), (36), (37) and Lemma 4, a straightforward computation leads to the following asymptotics

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \text{Crt}_{N,0}(t) \geq \frac{\gamma}{2} \log \left(\frac{\xi_1''}{\xi_1'} \right) + \frac{1-\gamma}{2} \log \left(\frac{\xi_2''}{\xi_2'} \right) + \sup_{a \in [0,1]} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{I}_N(t, a), \quad (41)$$

where

$$\mathcal{I}_N(t, a) = \int_{-\infty}^{\min\{t, a^*\}} e^{-\frac{Nx^2}{2}} \prod_{i=1}^2 \mathbb{E}_G^{N_i} \left[\exp \left\{ \frac{N_i}{2} \left((\lambda_0^{N_i})^2 - \frac{2\xi_i''}{\alpha_i^2} \left(\lambda_0^{N_i} - \frac{a \xi_i' x}{\gamma_i (2\xi_i'')^{\frac{1}{2}}} \right)^2 \right) \right\} \right] dx$$

with

$$a^* = -\frac{1 + \sqrt{\gamma_*}}{(1-a) \min\{\frac{\xi_1'}{\gamma}, \frac{\xi_2'}{1-\gamma}\}}.$$

The asymptotics of $\mathcal{I}_N(t, a)$ for a and t fixed and N going to infinity can be derived by an application of Laplace-Varadhan's lemma [16, Theorem 4.5.1]. Indeed, the smallest eigenvalue of the GOE matrix satisfies a LDP with speed N and rate function

$$I(x) = \int_{\sqrt{2}}^{|x|} \sqrt{z^2 - 2} dz = \frac{1}{2} \left(|x| \sqrt{x^2 - 2} + \log 2 - 2 \log \left(|x| + \sqrt{x^2 - 2} \right) \right) \quad (42)$$

if $x \leq -\sqrt{2}$ and $I(x) = \infty$ for $x > -\sqrt{2}$ (see [13, Theorem 6.2]). Therefore,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{I}_N(t, a) = \sup_{\substack{x \leq \min\{t, a^*\} \\ y_1, y_2 \leq -\sqrt{2}}} \left[-\frac{x^2}{2} + \sum_{i=1}^2 \gamma_i \left(\frac{y_i^2}{2} - \frac{\xi_i''}{\alpha_i^2} \left(y_i - \frac{a \xi_i' x}{\gamma_i (2\xi_i'')^{\frac{1}{2}}} \right)^2 - I(y_i) \right) \right].$$

Denoting the right-hand side of (41) as $J(t)$ ends the proof of the lower bound in the case $\alpha_1, \alpha_2 \neq 0$. To see that $J(t)$ is positive for some t , it suffices to take $y_1 = y_2 = -\sqrt{2}$ and $t \geq a^* = x$ in the equation above. This computation is similar to [5, Equation (2.15)]. The case when one (or both) $\alpha_i = 0$ is simpler as one (or both) of the parties becomes simply a pure p -spin model and one just needs to apply Lemma 5 with $Y = H_N$ as in [6, Equation (3.25)]. We leave the details to the reader.

References

- [1] R. J. Adler and J. E. Taylor (2007) *Random Fields and Geometry*. Springer Science.
- [2] D. Amit (1989) *Modeling Brain Function*. Cambridge University Press.
- [3] G. Anderson, A. Guionnet and O. Zeitouni (2010) *An introduction to random matrices*. Cambridge studies in advanced mathematics, **118**.
- [4] E. Agliari, A. Barra, A. Galluzzi, F. Guerra and F. Moauro (2012) Multitasking associative networks. *Phys. Rev. Lett.*, **109**, 268101.
- [5] A. Auffinger and G. Ben Arous (2013) Complexity of random smooth functions on the high-dimensional sphere. *Ann. of Probab.*, **41**, no. 6, 4214-4247.
- [6] A. Auffinger, G. Ben Arous, and J. Cerny (2013) Complexity of spin glasses and random matrices. *Comm. on Pure and Applied Math.*, **66**, 165-201.
- [7] J. Azais, M. Wschebor (2009) *Level Sets and Extrema of Random Processes and Fields*. Wiley Publications.
- [8] A. Barra, E. Agliari (2010) A statistical mechanics approach to autopoietic immune networks. *J. Stat. Mech.*, **07**, 07004.
- [9] A. Barra, P. Contucci (2010) Toward a quantitative approach to migrants integration. *Euro. Phys. Lett.*, **89**, 68001.
- [10] A. Barra, G. Genovese, F. Guerra and D. Tantari (2012) How glassy are neural networks? *J. Stat. Mech.*, P07009.
- [11] A. Barra, A. Galluzzi, F. Guerra, A. Pizzoferrato and D. Tantari (2013) Mean field bipartite spin models treated with mechanical techniques. arXiv:1310.5901.
- [12] A. Barra, G. Genovese and F. Guerra (2011) Equilibrium statistical mechanics of bipartite spin systems. *J. Phys. A: Math. and Theor.*, **44**, 245002.
- [13] G. Ben Arous, A. Dembo and A. Guionnet (2001) Aging of spherical spin glasses. *Probab. Theory and Rel. Fields*, **120**, 1–67.
- [14] D. S. Bernstein (2005) *Matrix Mathematics*. Princeton University Press.
- [15] W.-K. Chen (2013) The Aizenman-Sims-Starr scheme and Parisi formula for mixed p-spin spherical models. *Electron. J. Probab.*, **18**, no. 94, 1-14.
- [16] A. Dembo and O. Zeitouni (2010) *Large deviations techniques and applications*. Stochastic Modelling and Applied Probability, Springer-Verlag.
- [17] P.S. Dodds, R. Muhamad and D.J. Watts (2003) An Experimental Study of Search in Global Social Networks. *Science*, **301**, 5634.

- [18] Y. Fyodorov (2013) High-Dimensional Random Fields and Random Matrix Theory. arXiv:1307.2379
- [19] C. Martelli, A. De Martino, E. Marinari, M. Marsili and I. Perez Castillo (2009) Identifying essential genes in E. coli from a metabolic optimization principle. *Proc. Natl. Acad. Sc.*, **106**, 2607.
- [20] D. Panchenko (2009) Cavity method in the spherical SK model. *Ann. Inst. Henri Poincar Probab. Stat.*, **45**, no. 4, 1020-1047.
- [21] G. Parisi (1990) A simple model for the immune network. *Proc. Nat. Acad. Sc.*, **87**, 2412-2416.
- [22] M. Talagrand (2006) Free energy of the spherical mean field model. *Prob. Theory and Rel. Fields*, **134**, 339-382.
- [23] M. Talagrand (2011) *Mean field models for spin glasses*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, **54**, Springer-Verlag.